

BEST AVAILABLE COPY

DERWENT-ACC-NO: 2000-165383

DERWENT-WEEK: 200308

COPYRIGHT 2005 DERWENT INFORMATION LTD

TITLE: Test data production system for data
warehouse -
data production involves generating test data when
parameter into data starting unit input indication and
information and data production unit, based on schema
information generation rule and coupling

PATENT-ASSIGNEE: NEC SOFTWARE KANSAI LTD[NIDE]

PRIORITY-DATA: 1998JP-0191848 (July 7, 1998)

PATENT-FAMILY:

PUB-NO	PAGES	PUB-DATE	
LANGUAGE		MAIN-IPC	
JP 3366257 B2		January 14, 2003	N/A
007	G06F 017/30		
JP 2000020529 A		January 21, 2000	N/A
007	G06F 017/30		

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO
APPL-DATE		
JP 3366257B2	N/A	1998JP-
0191848	July 7, 1998	
JP 3366257B2	Previous Publ.	JP2000020529
N/A		
JP2000020529A	N/A	1998JP-
0191848	July 7, 1998	

INT-CL (IPC): G06F017/30

ABSTRACTED-PUB-NO: JP2000020529A

BASIC-ABSTRACT:

NOVELTY - A data production information memory (4) holds information about data generation rule and coupling information. An accessing unit (9) acquires schema information from database (8). A data production starting unit (6) input an indication and a parameter into data production unit (5) to generate test data based on schema information and other relevant information. DETAILED DESCRIPTION - A data memory (7) stores test data and outputs to database (8) via access unit. A data production information register (3) input data generation rule of test data and coupling information on an objective table. Test data is produced based on generation rule, coupling information and schema information.

USE - For testing database of data warehouse.

ADVANTAGE - Since database has data generation rule and coupling information for a schema information, performs automatic generation of test data. Since data production is performed by correction of schema information, increases efficiency of test data production operation. DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of test data manufacturing apparatus. (3) Data production information register; (4) Data production information memory; (5) Data production unit; (6) Data production starting unit; (7) Data memory; (8) Database; (9) Accessing unit.

CHOSEN-DRAWING: Dwg.1/8

TITLE-TERMS: TEST DATA PRODUCE SYSTEM DATA WAREHOUSE
GENERATE TEST DATA DATA
PRODUCE START UNIT INPUT INDICATE PARAMETER
DATA PRODUCE UNIT BASED
INFORMATION DATA GENERATE RULE COUPLE
INFORMATION

DERWENT-CLASS: T01

EPI-CODES: T01-G07A; T01-J05B4;

SECONDARY-ACC-NO:

Non-CPI Secondary Accession Numbers: N2000-123881

【特許請求の範囲】

【請求項1】 データウェアハウスのデータベースをテストするテストデータを作成するデータウェアハウステストデータ作成装置において、

データ生成ルール及び対象表の結合情報並びに前記データベースのスキーマ情報に基づいて、テストデータを生成することを特徴とするデータウェアハウステストデータ作成装置。

【請求項2】 前記データウェアハウステストデータ作成装置は、

前記データベースに入出力するデータベースアクセス手段と、

テストデータのデータ生成ルール及び前記対象表の結合情報を入力するデータ作成情報登録手段と、

前記データベースから前記データベースアクセス手段を通じて読み出した前記スキーマ情報と、前記データ生成ルール及び前記結合情報とで、なるデータ作成情報を保持するデータ作成情報記憶手段と、

前記スキーマ情報並びに前記データ生成ルール及び前記結合情報によりテストデータを生成するデータ作成手段と、

前記データ作成手段に指示及びパラメータを入力するデータ作成起動手段と、

前記テストデータ生成の作業領域であるデータメモリと、

を具備し、

前記データメモリのテストデータを前記データベースアクセス手段によって、前記データベースに出力することを特徴とするデータウェアハウステストデータ作成装置。

【請求項3】 前記スキーマ情報は、

表名と、列名と、属性と、桁数と、制約条件と、を有する定義情報を含むことを特徴とする請求項2記載のデータウェアハウステストデータ作成装置。

【請求項4】 前記データ生成ルールは、

データ生成方法の指定と、値の最大値と最小値の指定と、を含むことを特徴とする請求項2記載のデータウェアハウステストデータ作成装置。

【請求項5】 前記結合情報は、

結合する表の結合条件である、結合元表名と、結合元列名と、結合先表名と、結合先列名と、カーディナリティと、を含むことを特徴とする請求項2記載のデータウェアハウステストデータ作成装置。

【請求項6】 前記データウェアハウステストデータ作成装置を動作させる実行手段のプログラムを記憶する計算機読み取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、データウェアハウスのデータベースをテストするデータを作成すること

に関し、特にデータベースのスキーマ情報と、データ生成ルール及び結合情報とに基づいて、テストデータを自動的に生成するデータウェアハウステストデータ作成装置に関する。

【0002】

【従来の技術】大規模データウェアハウスを構築する際、データベースの検索性能に影響を与える1つの要因としスキーマ構造があげられる。このスキーマ構造が適切に設計されているか否かを検証するには、実際に構築したデータベースにサンプルデータを登録して性能検査を行う必要がある。このサンプルデータ作成は、実際の業務データを、構築したスキーマ構成に適合した形に加工して登録するか、作成するサンプルデータが大量になった場合、データ作成用のプログラムによって、データを作成している。

【0003】従来、スキーマ情報をデータ登録やデータ検索に利用することは行われている。特開平5-28154号公報によれば、文書を構成する情報要素及びその階層関係を定義する文書構造情報に基づく入力画面と、文書構造情報に基づくスキーマをデータベースに構築して、文書をデータベースに容易に登録することが開示されている。利用者はスキーマを作成しておくだけで、該スキーマに基づく入力画面によって、文書登録を容易に行うことができる。また、特開平6-223118号公報によれば、データベースの論理定義情報を保持し、該定義情報によりデータ検索することが開示されている。アプリケーションプログラムに対して論理的なデータベースが該定義情報で定義されているので、検索対象が複数のテーブルに格納されている状態のデータ検索が容易に行える。

【0004】この従来技術では、入力したデータをデータベースに登録すること、あるいは外部からのデータベース検索を容易にすることにスキーマ情報を利用している。データベースのテストや評価に使用するテストデータ作成ため、上述の業務データの加工やテストデータ作成プログラムに代えてスキーマ情報を利用することの従来例は見あたらない。

【0005】

【発明が解決しようとする課題】第1の問題点は、スキーマ情報が変更になるたびに新しいスキーマ情報に合わせてデータ作成プログラムを変更しなければならないことである。

【0006】第2の問題点は、データを生成する際のパターンをプログラム中に保持しているため、パターンの異なるデータを生成するためには該プログラムを変更しなければならないことである。

【0007】この発明は、データベースが有するスキーマ情報に、データ生成ルールと結合情報とを加えて、テストデータを自動生成するデータウェアハウステストデータ作成装置にある。

【0008】

【課題を解決するための手段】そのため、この発明のデータウェアハウスのデータベースをテストするテストデータを作成するデータウェアハウステストデータ作成装置において、データ生成ルール及び対象表の結合情報並びに前記データベースのスキーマ情報に基づいて、テストデータを生成することを特徴とする。

【0009】更に、前記データウェアハウステストデータ作成装置は、前記データベースに入出力するデータベースアクセス手段と、テストデータのデータ生成ルール及び前記対象表の結合情報を入力するデータ作成情報登録手段と、前記データベースから前記データベースアクセス手段を通じて読み出した前記スキーマ情報と、前記データ生成ルール及び前記結合情報とで、なるデータ作成情報を保持するデータ作成情報記憶手段と、前記スキーマ情報並びに前記データ生成ルール及び前記結合情報によりテストデータを生成するデータ作成手段と、前記データ作成手段に指示及びパラメータを入力するデータ作成起動手段と、前記テストデータ生成の作業領域であるデータメモリと、を具備し、前記データメモリのテストデータを前記データベースアクセス手段によって、前記データベースに出力することを特徴とする。

【0010】

【発明の実施の形態】この発明について、図面を参照して説明する。この発明の第1の実施の形態を示す図1を参照すると、データウェアハウステストデータ作成装置1は、テストデータ作成に必要なデータ生成ルールと結合情報とを入力するデータ作成情報登録手段3と、データベースのスキーマ情報とデータ生成ルールと結合情報とでなるデータ作成情報を保持するデータ作成情報記憶手段4と、該データ作成情報に基づいてテストデータを作成するデータ作成手段5と、入力画面から指示及びパラメータを入力してデータ作成手段5を起動するデータ作成起動手段6と、テストデータを生成する作業領域のデータメモリ7と、データウェアハウスが有するデータベース8と、データベース8に入出力するデータベースアクセス手段9と、を備える。

【0011】更に、この実施の形態の細部を示す図2を参照すると、データ作成情報登録手段3は、データベース8から読み込んだスキーマ情報に対して、テストデータ作成のデータ生成ルールを定義するデータ生成ルール登録手段30と、データベースのスキーマ間の結合情報を定義する結合情報登録手段31と、を有している。データ作成情報記憶手段4は、データベース8から読み込んだスキーマ情報及びデータ作成に必要な生成ルール等を格納するスキーマ情報ファイル40と、データベースのスキーマ間の結合情報を格納する結合情報ファイル41と、を有している。

【0012】データ作成手段5は、結合情報ファイル41に格納されている結合情報と、スキーマ情報ファイル

40に格納されているスキーマ情報及びデータ生成ルールを読み込む作成情報読み取り手段50と、結合元表であるファクトテーブルに結合される結合先表であるマスタテーブルの結合キーを読み込んで結合キーテーブルを作成する結合キーテーブル作成手段51と、スキーマ情報及びデータ生成ルール並びに結合キーテーブルをもとにデータ作成を行うデータ生成手段52と、作成したテストデータをデータメモリ7に書き込むデータ出力手段53と、を備えている。データベースアクセス手段9は、データベース8のスキーマ情報を読みとるスキーマ情報取り込み手段90と、データベース8のデータを読み込んでデータメモリ7に送出するデータ読み込み手段91と、データメモリ7の内容をデータベース8に登録するためのデータ書き込み手段92と、を有している。

【0013】スキーマ情報ファイル40の構造を示す図3を参照すると、スキーマ情報ファイル40は、データベース8からスキーマ情報取り込み手段90によって取り込まれた表名311と、列名312と、属性313と、桁数314と、制約315と、に対して、データ生成ルール登録手段30を通じて入力されたデータ生成ルール316を付加して構成される。項目のうち、表名311、列名312、属性313、桁数314、制約315はデータベース8が格納しているスキーマの定義情報である。制約315には、重複の有無やNULL値を認めるかどうか格納される。データ生成ルール316は、データ生成ルール登録手段30によって、外部設定される。

【0014】結合情報ファイル41の構造を示す図4を参照すると、結合情報ファイル41は、結合情報登録手段31を通じて入力される結合元表名411と、結合元列名412と、結合先表名413と、結合先列名414と、カーディナリティ415と、から構成される。結合元列名412及び結合先列名414は、表を結合させる際のキーとなる列名である。カーディナリティ415には、1:1、1:n、n:1、n:n等の結合の対応関係が格納される。

【0015】この第1の実施の形態の各手段の機能細部について、図2を参照して説明すると、スキーマ情報取り込み手段90は、データベース8中の各テーブル毎のスキーマ情報を取得する。取得するスキーマ情報の項目は、表名311と、列名312と、属性313と、桁数314と、制約315と、でなる。これらの取得した項目は、スキーマ情報ファイル40に登録する。スキーマ情報ファイル40に登録される内容を例示する図3を参照すると、表名311に、Fact1、Master1、Master2の3つの表があり、行301乃至行305は、Fact1の各列、行306乃至行308は、Master1の各列、行309、行310は、Master2の各列の項目である。結合情報登録手段31から各スキーマ間の結合情報を入力し、結合情報ファ

イル41に登録する。結合情報ファイル41に登録される内容を例示する図4を参照すると、行401では、結合元表名411のFact1と結合先表名413のMaster1とが、結合元列名Fact1col3と結合先列名Master1col1とが結合し、そのカーディナリティはN:1である。また、行402では、結合元表名411のFact1と結合先表名のMaster2とが、結合元列名Fact1col4と結合先列名Master2col1とが結合し、そのカーディナリティはN:1である。次に、データ生成ルール登録手段30では、データ生成時に必要なデータ生成ルールを入力し、スキーマ情報ファイル40に格納する。図3では、データ生成ルール316が、各行301乃至行310に格納されている。図5に例示する書式500を用いて、データ生成ルール316の内容を説明すると、データ生成ルール316はキーワードパラメータの形式で指定される。「MODE」には、データを生成する方法を指定する「R」(RANDOM)と、「I」(INCREASE)と、「W」(WAVE)と、「F」(FIX)と、「P」(PREFIX)と、「M」(MASTER)と、の6種類の内のいずれかが指定される。「FROM」および「TO」パラメータは「R」、「I」、「W」の指定の際に有効となり、値の最小値と最大値とを指定する。「MARGIN」パラメータは「I」、「W」の指定の際に有効となり、各値のデータ間隔を表す。

【0016】上記のデータ生成ルールを適用して、生成されるデータイメージを図6に例示する。「RANDOM」指定601には、「FROM」で指定された値から「TO」で指定された値までの範囲内でランダムな数値を生成する。この例では、「1」から「100」までの値がランダムに生成される。次に「INCREASE」指定611には、「FROM」で指定された値から、「TO」で指定された値まで「MARGIN」パラメータで指定された値を増分にした数値が生成される。「TO」が指定されていない場合は、作成時に指定した件数まで作成する。この例では、「1」から「100」まで「2」を増分にした値が生成され、「1, 3, 5, 7」の数値が生成される。「WAVE」指定621には、「FROM」で指定された値から、「TO」で指定された値まで「MARGIN」パラメータで指定された値で増減を繰り返す。この例では、「1」から「100」まで「1」ずつ増加した値を生成した後、「100」に達した後は、「1」ずつ減少した値を生成する。「FIX」指定631には、固定の値を生成するが、複数指定された場合はその値の中からランダムに選択される。この例では、「FIX」値として、「aaaaa」と、「bbbbbb」と、の2つの値が指定されているため、この2つの値をランダムに選択して生成する。「PREFIX」指定641には「PREFIX」指定値にキー

値を加えた値を選択する。複数指定された場合はその中からランダムに値が選択される。この例では、プリフィックス値「aaaaa」にキー値をつけた値が生成される。「MASTER」指定651には、結合対象となるマスタテーブルの結合キーを値とし、ランダムに生成する。この例では、結合キーテーブル660に結合先列のキー値「101」乃至キー値「109」が登録されているため、この中からランダムに選択された値を生成する。

【0017】次に、データ作成起動手段6を用いて、作成対象となる表を選択し、作成するデータ件数等のパラメータを指定し、データ作成手段5を起動する。作成情報読み取り手段50は、作成対象となる表に関するスキーマ情報ファイル40及び結合情報ファイル41を取得する。結合キーテーブル作成手段51は、結合対象の表のデータをデータベース8からデータベースアクセス手段9を通じて取得し、対象の表の結合キーとなる列名のデータをデータメモリ7上に作成する。データ生成手段52は作成情報読み取り手段50で読み込んだスキーマ情報ファイル40及び結合情報ファイル41に基づいて、対象となる表データの生成をデータ生成ルール316によって行い、データ出力手段53に引き渡す。データ出力手段53は、データメモリ7にデータ出力を行う。データメモリ7のデータはデータ書き込み手段92によってデータベース8に書き込まれる。

【0018】この第1の実施の形態の動作を説明する。データ作成手段5の処理手順を表すフローの図7を参照すると、まず初期設定を行う(ステップ701)。データ作成の対象となる表名及び作成データ件数等の作成パラメータをデータ作成起動手段6を通じて取得し(ステップ702)、スキーマ情報ファイル40を読み込み、その表のスキーマ情報を取得する(ステップ703)。次に、作成対象となるテーブルのテーブル種別を判断する(ステップ704)。テーブル種別が結合元表のファクトテーブルの場合(ステップ704のYes)、該ファクトテーブルに結合される結合先表のマスタテーブルの結合先列名414及びカーディナリティ415を結合情報ファイル41より取得する(ステップ705)。取得したマスタテーブルと結合列の情報を読み込み対象マスタテーブル情報としてセットする(ステップ706)。セットされた対象マスタテーブルのデータを全て読み込み、結合キーテーブル660としてデータメモリ7に登録する(ステップ707)。結合キーテーブル601未作成のマスタテーブルが存在するかどうかを判断し(ステップ708)、結合キーテーブル未作成のマスタテーブルが存在する場合(ステップ708のYes)、ステップ706に戻る。結合キーテーブル未作成のマスタテーブルが存在しない場合(ステップ708のNo)、データ生成ルールの解釈を行う(ステップ709)。解釈したデータ生成ルール及び結合キーテーブル

の情報に基づいて、データ生成を行い(ステップ710)、データ出力を行う(ステップ711)。データ作成終了したか否かを判断し(ステップ712)、データ作成終了の場合(ステップ712のYes)、終了する。データ作成終了でない場合(ステップ712のNo)、ステップ710に戻る。ステップ704の判定で、テーブル種別がファクトテーブルでない場合(ステップ708のNo)、ステップ709に進む。

【0019】次に、この発明の第2の実施の形態について図面を参照して説明する。第1の形態の細部構成を示す図2に対して、データ作成情報登録手段3にスキーマ情報登録手段32と、データベースアクセス手段9にスキーマ情報登録手段93とを付加する変形を施した第2の実施の形態のデータウェアハウステストデータ作成装置2を図8に示す。スキーマ情報登録部93は、スキーマ情報ファイル40の表名311、列名312、属性313、桁数314、制約315等の定義情報をデータベース8に登録し、データベース8上にスキーマを構築する。このデータウェアハウステストデータ作成装置2では、このデータウェアハウステストデータ作成装置1をデータベース8のスキーマ定義装置として拡張することで、スキーマ変更時の作業をさらに効率化できるという新たな効果を有する。

【0020】

【発明の効果】第1の効果は、データベースが有するスキーマ情報に基づいて、テストデータ生成を行うため、スキーマ変更の度にテストデータ作成プログラムを変更あるいは作成する必要がなく、スキーマ情報の修正だけでデータ作成が行えるためテストデータ作成作業を効率化できるということである。

【0021】第2の効果は、テストデータ作成時に対象となる表及び項目にデータ生成ルール及び結合情報を指定することができるため、この生成ルールを変更するだけで、対象表に対して、様々なパターンでのテストデータ作成が行えるということである。

【図面の簡単な説明】

【図1】この発明の第1の実施の形態を示す図である。

【図2】図1の細部を示す図である。

【図3】図2のスキーマ情報ファイルの構成を例示する図である。

【図4】図2の結合情報ファイルの構成を例示する図である。

【図5】図2のデータ生成ルールを例示する図である。

【図6】生成されたデータを例示する図である。

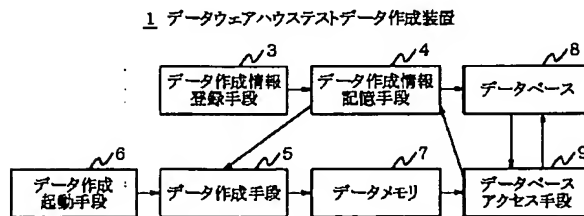
【図7】図2のデータ作成手段の動作を示す図である。

【図8】この発明の第2の実施の形態を示す図である。

【符号の説明】

- 1, 2 データウェアハウステストデータ作成装置
- 3 データ作成情報登録手段
- 4 データ作成情報記憶手段
- 5 データ作成手段
- 6 データ作成起動手段
- 7 データメモリ
- 8 データベース
- 9 データベースアクセス手段
- 40 スキーマ情報ファイル
- 41 結合情報ファイル
- 311乃至315 スキーマの定義情報
- 315 データ生成ルール

【図1】



【図4】

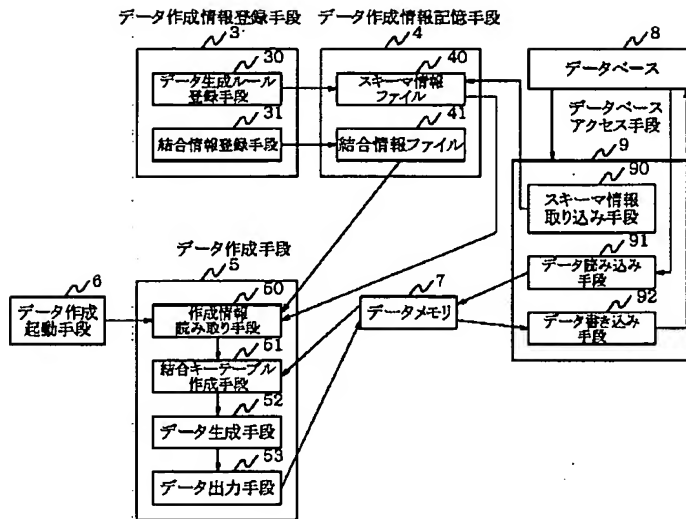
41 結合情報ファイル

	411 結合元表名	412 結合元列名	413 結合先表名	414 結合先列名	415 カーディナリティ
401	Fact1	Fact1col3	Master1	Master1col1	N:1
402	Fact1	Fact1col4	Master2	Master2col1	N:1

【図5】

- 600 書式
- 601 RANDOM指定
- 602 Mode=R from=xxxxx to=xxxxx
- 603 INCREASE指定
- 604 Mode=I from=xxxxx (to=xxxxx) margin=xxx
- 605 WAVE指定
- 606 Mode=W from=xxxxx to=xxxxx margin=xxx
- 607 FIX指定
- 608 Mode=F Fixvalue1,Fixvalue2,....
- 609 PREFIX指定
- 610 Mode=P Prefix1,Prefix2,....
- 611 MASTER指定
- 612 Mode=M

【図2】



【図3】

40 スキーマ情報ファイル

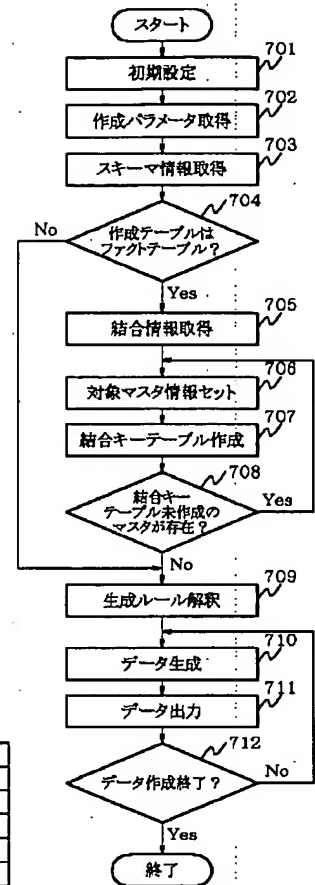
311	312	313	314	315	316
表名	列名	属性	桁数	制約	データ生成ルール
Fact1	Fact1col1	Num	5	NotNull,unique	Mode=I from=1 margin=1
Fact1	Fact1col2	Date	8		Mode=R from=19970101 to=19980331
Fact1	Fact1col3	Num	6	NotNull	Mode=M
Fact1	Fact1col4	Num	4	NotNull	Mode=M
Fact1	Fact1col5	Char	16		Mode=P Prefix1,Prefix2
Master1	Master1col1	Num	6	NotNull,unique	Mode=I from=1 margin=1
Master1	Master1col2	Char	20		Mode=P Mastername
Master1	Master1col3	Char	10		Mode=F Fixvalue
Master2	Master2col1	Num	4	NotNull,unique	Mode=I from=1 margin=1
Master2	Master2col2	Char	20		Mode=P Mastername

【図6】

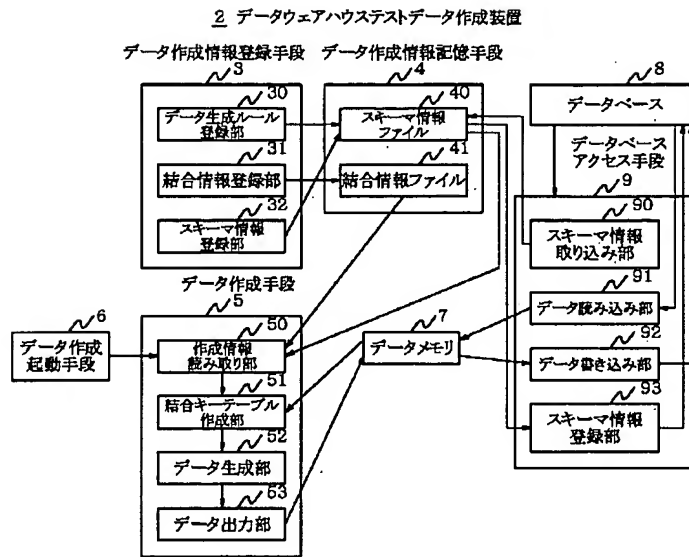
601 ~RANDOM指定
 602 ~Mode=R from=1 to=100
 603 ~3, 8, 94, 55, 32, 13, 98, ...
 611 ~INCEREAASE指定
 612 ~Mode=I from=1 to=100 margin=2
 613 ~1, 3, 5, 7, 9, 11, 13, 15, 17, 19, ... 97, 99
 621 ~WAVE指定
 622 ~Mode=W from=1 to=100 margin=1
 623 ~1, 2, 3, 4, ... 98, 99, 100, 99, 98, ... 3, 2, 1, 2...
 631 ~FIX指定
 632 ~Mode=F aaaaa,bbbbbb
 633 ~aaaaa,bbbbbb,bbbbbb,aaaaa,bbbbbb, ...
 641 ~PREFIX指定
 642 ~Mode=P aaaaa
 643 ~aaaaa1,aaaaa2,aaaaa3, ...
 651 ~MASTER指定
 652 ~Mode=M
 653 ~101, 106, 102, 101, 105

660 結合キーテーブル
 661 ~101 102 103 104 105 106 107 108 109

【図7】



【図8】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.